U.S. Department of the Interior
U.S. Geological Survey

The National Map Catalog Technical Discussion Paper

Availability and Growth of *The National Map* in Fiscal Year 2005

November 2005

Larry Moore
Marshall Creighton

NGTOC III

1 The Catalog Service Checker

The need to quantify Web Map Service (WMS) availability was recognized early in the development of *The National Map* system. In July 2003 a software "service checker" was deployed to monitor the availability of services that contribute data to *The National Map*. Data from this monitoring is stored in Oracle tables in the catalog database.

Service **availability** and **reliability** are synonyms in this document, and are defined for each service as

$$\text{availability} = \frac{\text{'available' data points}}{\text{total data points}}$$

The data used to calculate this value are not continuous; each service is queried about every 17 minutes. The service checker itself is not 100% reliable, so there are some gaps in the data. However, for time spans of weeks enough data points are collected to approximate continuous monitoring.

The basic unit of availability used in this report is one **service-month**, calculated with the above equation for all sample points for one service for one month.

The scope of this document is the availability of individual WMSs that contribute data to *The*

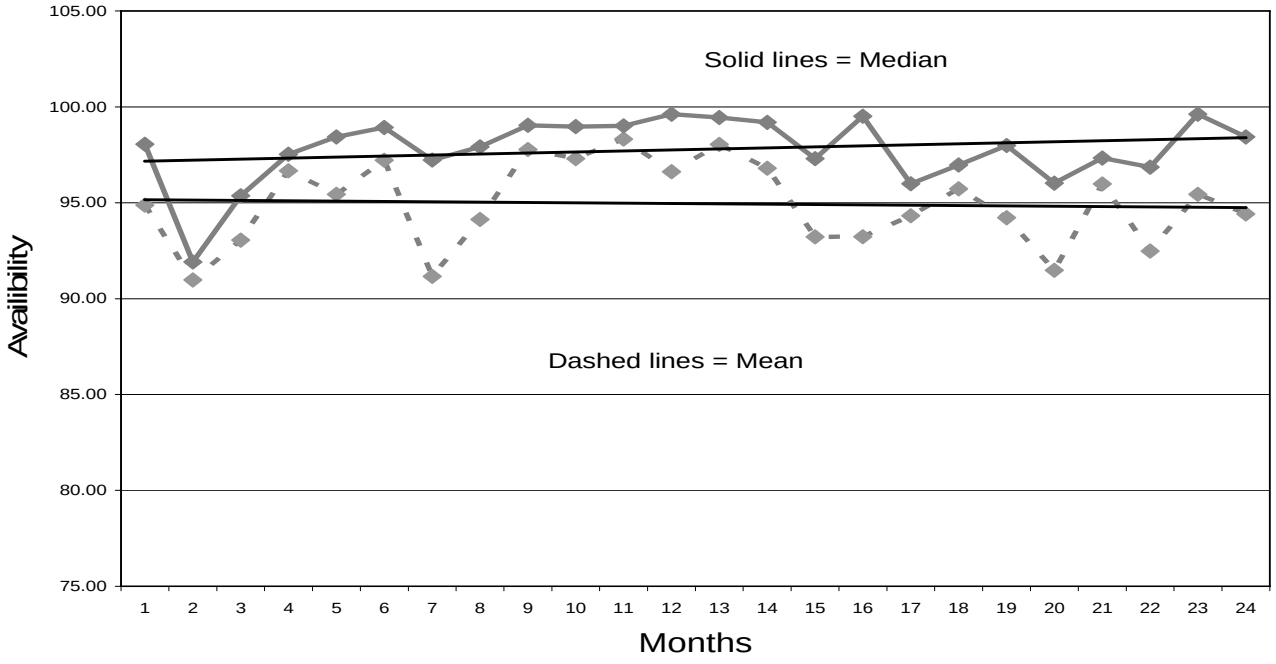


Figure 1. Service availability for FY04 - FY05. The dashed gray line connects the arithmetic **mean** of all service availability values for each month. The solid gray line connects the **median** of all service availability values for each month. The black lines are linear regression lines. Mean values are lower than median values because it is common for a service to have short periods of extreme unreliability, especially early in the life of the service. These periods drag the mean down, but have little effect on the median. Which statistic is a better measure of reliability is a matter of taste and priorities.

National Map. Overall reliability of *The National Map* system is a much larger issue, affected by many additional factors.

See section 4 for references to documents that contain more information about the operation of the service checker.

2 FY05 Data Summary

2.1 Service Availability

About 170 services contributed data to *The National Map* applications at some time in FY05. Figure 1 illustrates overall service availability by month for FY04 and FY05

The following table is a different view of the data used to create figure 1. The table lists services that were registered in the catalog and provided data to either the public viewer or the SOLDGR application in FY05. It shows more detail per service than figure 1, but less detail per month. The list is sorted first by the number of months each service was registered with *The National Map*, and second by median availability. The third column is the median of the monthly availability statistics for the service. Medians tend to be higher than means, for the same reasons the median line is above the mean line in figure 1.

Services at the top of the list were available a high percentage of the time all year. Services at the bottom of the list were part of *The National Map* for fewer months, or were not as reliable, or both. Note also that if the table were sorted on the mean column rather than the median column, the order would change slightly.

SERVICE NAME	Mean avail %	Median avail %	Months online
USGS Catalog WMS	100.0	100.0	12
TerraServer USA	99.8	99.9	12
NEXRad WMS	99.5	99.9	12
Minnesota DOT MapServer	98.0	99.9	12
York County (SC) WMS (NC OneMap)	97.9	99.8	12
Seattle-Tacoma, WA WMS	99.2	99.8	12
Kansas (RNMP-DASC) WMS	99.1	99.8	12
Wash-ID Pilot WMS	99.6	99.8	12
Sedgwick County (KS) WMS	99.7	99.8	12
Tahoe Pilot WMS	98.1	99.7	12
Sacramento, CA WMS	98.1	99.7	12
Arkansas (RNMP-CAST) WMS	91.8	99.7	12
MSDIS WMS	96.0	99.6	12
Montana WMS	98.8	99.5	12
USGS Wetlands WMS	98.9	99.5	12
USGS Wetlands Alaska	98.4	99.4	12
St. Louis County (MO) WMS	99.0	99.4	12
US Forest Service Southwestern Regi	97.8	99.3	12
US Forest Service Pike NF Pilot	97.8	99.3	12
US Forest Service Bankhead NF	97.6	99.3	12
US Forest Service Boundaries WMS	97.8	99.3	12
NHDGeo WMS (RMMC)	98.1	99.2	12
USGS Ref WMS (RMMC)	98.5	99.2	12
Minnesota LMIC Health	97.2	99.1	12
Bozeman /Gallatin MT	78.4	99.0	12
Minnesota LMIC Boundaries	97.0	99.0	12
Minnesota LMIC Geology	97.0	99.0	12
Minnesota LMIC Social	97.2	99.0	12
Missoula County MT Geographic Commu	93.7	99.0	12
Oklahoma WMS Server (GEO)	98.6	99.0	12
Texas TNRIS DOQQs UTM Zone14	97.4	98.9	12
Texas TNRIS DOQQs UTM Zone13	97.3	98.9	12
Utah Pilot WMS	98.0	98.8	12
National Park Service	98.0	98.8	12
Denver CO Pilot WMS	98.0	98.8	12
Texas TNRIS DOQQs UTM Zone15	97.2	98.8	12
Albuquerque NM Pilot WMS	95.8	98.8	12
Rocky's 20 (Backup Shaded Relief)	98.0	98.8	12

SERVICE NAME	Mean avail %	Median avail %	Months online
W			
Rocky Mountain National Park Pilot	98.0	98.7	12
Lubrecht(UMCFC-Montana)	87.2	98.7	12
National Atlas Data	96.3	98.5	12
RMMC Hazards Service (OGCConnector)	97.8	98.4	12
NC OneMap WMS	90.3	98.4	12
National Map Corp 2 WMS	94.7	98.1	12
NDEP WMS	93.1	97.9	12
US Forest Service Roads	96.0	97.6	12
USGS EDC LandCover MODIS 7day	95.2	96.4	12
USGS LANDSAT7 (EDC) WMS	95.3	96.4	12
USGS REF WMS (EDC)	95.3	96.4	12
USGS Shuttle Radar Topography Missi	95.3	96.4	12
NC OneMap Disaster_Response_Data	95.3	96.4	12
Mecklenburg Pilot (Mick_Co) WMS (NC	95.2	96.4	12
USGS BTS Roads WMS (EDC)	95.3	96.4	12
USGS Greenness (NDVI) (TEST)	95.3	96.4	12
New Jersey 2002 Hi-Res Ortho	95.3	96.4	12
USGS GTOPO WMS (EDC)	94.2	94.9	12
Texas TNRIS 2 WMS	90.1	94.4	12
Oahu, HI WMS	78.8	90.6	12
Kauai, HI WMS	78.8	90.6	12
GNIS/Atlas County Boundaries WMS	88.4	90.6	12
Molokai, HI WMS	78.8	90.5	12
GNIS WMS	88.9	90.5	12
Charleston County Roads	88.1	90.4	12
Shenandoah National Park	88.0	90.3	12
GNIS/Atlas State Boundaries WMS	88.1	90.2	12
Shenandoah National Park 1 meter im	88.0	90.2	12
Charleston County 1 Meter Orthos	88.3	89.7	12
Loudoun County (VA) WMS	88.3	89.7	12
Charleston 1-foot Resolution Imager	88.5	89.6	12
Washington DC Roads WMS	89.6	89.6	12
Washington DC (CUES DC Orthos) WMS	89.6	89.5	12
Merged VDOT, BTS, and CENSUS Roads	89.0	88.9	12
Delaware WMS	60.8	77.9	12
MetroGIS WMS (MN)	77.4	77.6	12
USGS Map Series Indices	97.1	99.1	11
Iowa Transportation	96.6	99.0	11
WORLDVUE	93.7	98.2	11

SERVICE NAME	Mean avail %	Median avail %	Months online
WRD Studies	93.7	98.2	11
NTAD	95.5	98.2	11
90TH RRC	94.1	97.9	11
USGS NED WMS (EDC)	95.7	97.9	11
Missouri Roads	95.2	97.8	11
USGS NLCD WMS (EDC)	95.7	97.8	11
Maui, HI WMS	76.9	88.6	11
Johnston County, NC	99.6	99.9	10
Kansas 2002 DOQQ	99.8	99.9	10
Flood Data (Harvard Design and Mapp	96.8	98.5	10
National Transportation Atlas Datab	95.0	98.1	10
NC OneMap Disaster Response Data 19	94.3	94.7	10
USGS EDC LandCover MODIS 2004	94.3	94.7	10
Fort Smith TNM Orthos	86.6	91.9	10
Wayne Co. WMS (IndianaMap)	99.1	100.0	9
USGS Wetlands Pacific Trust	98.8	99.9	9
USGS Wetlands Puerto Rico-US Virgin	98.8	99.9	9
Vermont NM Service	95.3	99.8	9
USGS Wetlands Hawai'i	97.9	99.3	9
Indiana Atlas (OGCConnector Test)	97.1	97.8	9
US Forest Service Culture	95.2	96.7	9
USGS NED WMS 1/9 Arc Second (EDC)	94.1	94.0	9
USGS NED WMS 1/3 Arc Second (EDC)	94.1	94.0	9
USGS EDC Ortho Urban WMS	94.2	94.0	9
Hawaii (Big Island), HI WMS	71.4	88.4	9
University of MO-Columbia WMS	99.7	99.9	8
Kentucky NM Service	99.3	99.9	8
NOAA's ENC Direct to GIS WMS	87.1	98.1	8
Afghanistan WMS	95.2	94.5	8
USGS EDC National Atlas	94.2	93.8	8
USGS EDC LandCover NLCD2001	93.4	93.8	8
USGS EDC Topobathy WMS	92.4	92.5	8
Fort Smith Arkansas	91.7	92.2	8
Iowa DNR National Map BU	99.7	100.0	7
GDT Dynamap 1000-US West	97.7	99.3	7
NDOP WMS	93.2	96.4	7
Denali National Park Satellite Imag	90.7	96.3	7
Africa WMS	96.0	96.1	7
Niihau, HI WMS	68.3	88.6	7
Lanai, HI WMS	68.5	88.6	7
Kahoolawe, HI WMS	68.4	88.6	7
Nebraska DNR New	99.3	99.7	6
GDT Dynamap 1000-US East	97.1	97.8	6

SERVICE NAME	Mean avail %	Median avail %	Months online
Geologic_map (OGCConnector v1.1)	86.7	97.4	6
National Map Corp WMS	89.2	97.4	6
Charleston County WMS	85.9	89.1	6
Arkansas DOQQ (CAST) WMS	86.0	84.2	6
Oklahoma Color DOQQ (SCI)	95.7	99.7	5
Oklahoma DRG (SCI)	95.7	99.7	5
Oklahoma DOQQ (SCI)	95.7	99.7	5
Kansas TNM GAP	95.6	99.4	5
world_volcanoes (OGCConnector v1.2)	95.8	98.4	5
Road Shields (USGS) (OGCConnector	95.7	98.1	5
Richland County SC WMS	99.9	99.9	4
Kansas TNM PowerPlants	98.0	99.8	4
Kansas TNM ElectricLines	98.0	99.7	4
Iowa Geographic Map Server	99.4	99.7	4
USFWS Habitat WMS	97.3	97.2	4
Benton_Co_AR_Orthoimagery	98.7	99.9	3
USGS EDC Elev SRTM	99.3	99.7	3
USGS EDC Ortho DOQQ	99.4	99.4	3
earth_weather (OGCConnector v1.2)	95.5	95.9	3
Washington DC WMS	79.6	76.1	3
Washington DC Hypso	78.6	74.3	3
TerraServer2 Map Server	100.0	100.0	2
NHD Labels	99.4	99.4	2
Washington_Co_AR_Orthoimagery	98.2	98.2	2
US Fish and Wildlife Service	98.1	98.1	2
National Map Corp3 WMS	97.7	97.7	2
North Dakota GIS Hub WMS - NDWMS_Ge	100.0	100.0	1
National Weather Service Hurricanes	99.9	99.9	1
Hurricane Katrina Data WMS	99.7	99.7	1
Indiana Marion County (Web Map Ser	99.7	99.7	1
Transportation Database Labels(CENS	99.0	99.0	1
Transportation Database (Census)	98.9	98.9	1
Katrina Photos (PRISM)	98.4	98.4	1
Disaster Response Data WMS	98.3	98.3	1
Katrina_storm_data (OGCConnector v1	98.1	98.1	1
Disaster Response Data 19 WMS	97.8	97.8	1
USGS katrina topomaps (OGCConnector	97.7	97.7	1
Oil and Gas Assessment - Yukon Fla	97.5	97.5	1
Oil and Gas Assessment - Western G	95.6	95.6	1

SERVICE NAME	Mean avail %	Median avail %	Months online
1 gulfcoast_miocene1 (OGCConnector	95.5	95.5	1
Oil and Gas Assessment - East Texa	95.2	95.2	1
Oil and Gas Assessment - Florida P	95.2	95.2	1
Oil and Gas Assessment - Southern	95.1	95.1	1
Oil and Gas Assessment - Northern	94.4	94.4	1
Charleston County WMS	85.9	89.1	6
Arkansas DOQQ (CAST) WMS	86.0	84.2	6
Oklahoma Color DOQQ (SCI)	95.7	99.7	5
Oklahoma DRG (SCI)	95.7	99.7	5
Oklahoma DOQQ (SCI)	95.7	99.7	5
Kansas TNM GAP	95.6	99.4	5
world_volcanoes (OGCConnector v1.2)	95.8	98.4	5
Road Shields (USGS) (OGCConnector	95.7	98.1	5
Richland County SC WMS	99.9	99.9	4
Kansas TNM PowerPlants	98.0	99.8	4
Kansas TNM ElectricLines	98.0	99.7	4
Iowa Geographic Map Server	99.4	99.7	4
USFWS Habitat WMS	97.3	97.2	4
Benton_Co_AR_Orthoimagery	98.7	99.9	3
USGS EDC Elev SRTM	99.3	99.7	3
USGS EDC Ortho DOQQ	99.4	99.4	3
earth_weather (OGCConnector v1.2)	95.5	95.9	3
Washington DC WMS	79.6	76.1	3
Washington DC Hypso	78.6	74.3	3
TerraServer2 Map Server	100.0	100.0	2
NHD Labels	99.4	99.4	2
Washington_Co_AR_Orthoimagery	98.2	98.2	2
US Fish and Wildlife Service	98.1	98.1	2
National Map Corp3 WMS	97.7	97.7	2
North Dakota GIS Hub WMS - NDWMS_Ge	100.0	100.0	1
National Weather Service Hurricanes	99.9	99.9	1
Hurricane Katrina Data WMS	99.7	99.7	1
Indiana Marion County (Web Map Ser	99.7	99.7	1
Transportation Database Labels(CENS	99.0	99.0	1

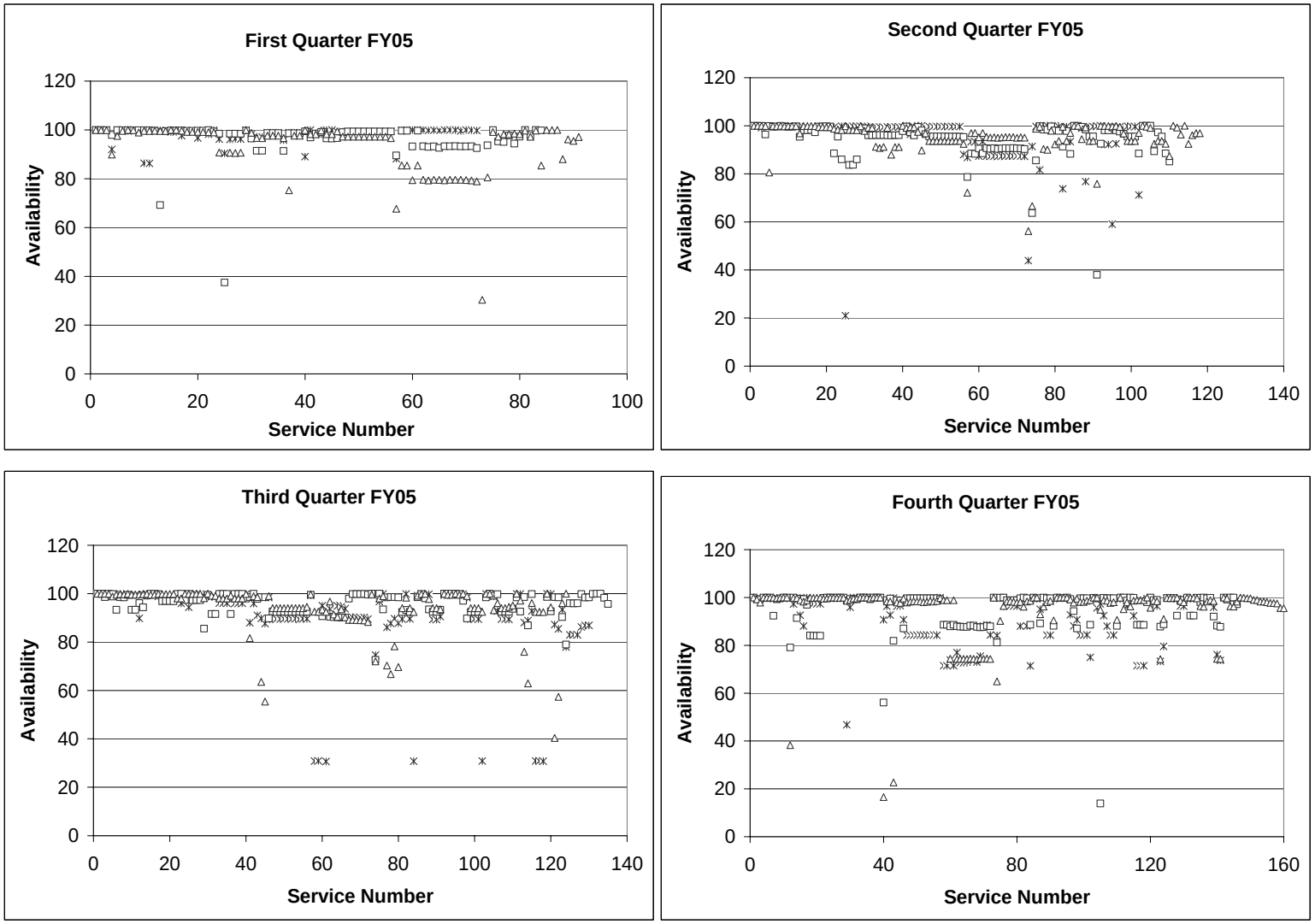


Figure 2. Scatter plots of service availability by quarter. Each point symbol represents one service-month. Within each quarter, stars are the first month, squares are the second, and triangles the third. Services (X axis) are not sorted in any particular order, but are in the same order for all four plots. The general pattern toward higher availability as more services are added is visually obvious.

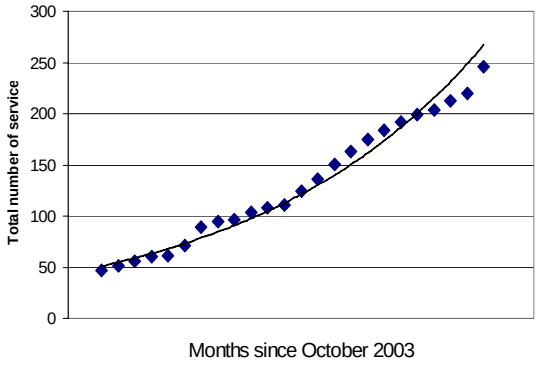
Figure 2 is a set of scatter plots of service-month availability points. Four plots are shown, one for each 3-month quarter. Each point on each plot represents one service-month. The graphs visually illustrate that as the year progressed:

- The number of services increased.
- Average reliability improved.
- Noticeable drops in service availability as providers improve services or add layers

2.2 Growth of The National Map Catalog

October 2003 - September 2005

*New **services** are being added to the database at an overall average rate of 9 per month.*

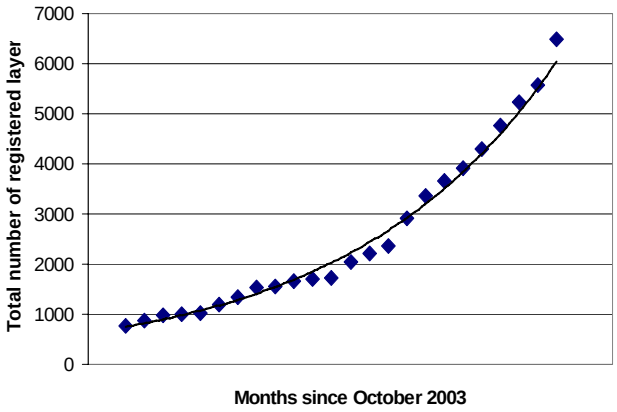


On both charts, the regression lines are best-fit exponential curves. The correlation coefficient R^2 is slightly above 0.98 in both cases. In both cases, the correlation for an exponential regression line is better than the correlation for a linear regression line. This means that not only is the database growing, but the **rate of growth** is also increasing.

Though the total number of layers exceeds 6000, only about 1800 are visible in the public viewer. Of the remainder, some are sensitive data in password protected services, made available through other applications; some are duplicative layers; some are data not within the scope of *The National Map*.

There are currently 114 partner organizations registered in the catalog.

*New **layers** are being added to the database at an overall average rate of 240 per month.*



2.3 Viewer Requests

A new monitoring capability was implemented in November 2004. We are now sampling and storing information about requests that come to the catalog service. This will not tell us who uses *The National Map* (from the catalog service perspective, the overwhelming majority of requests come from the public viewer, which does not pass along the identity of the "real" user), but it will tell us what data categories and spatial extents are being accessed most frequently.

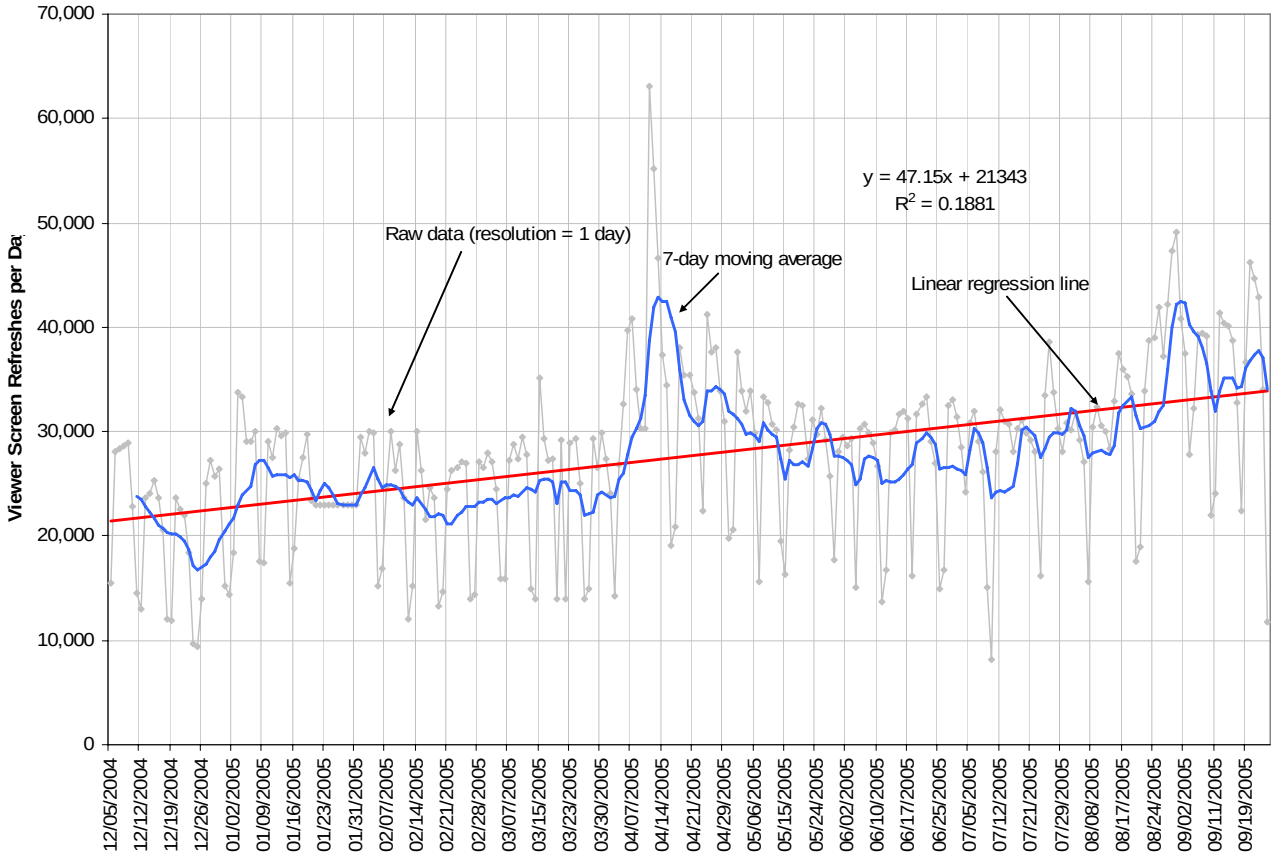


Figure 3. Number of viewer screen refresh requests

Figure 3 represents the number of viewer screen refreshes per day from December 5th, 2004 until September 30th, 2005. The see-saw pattern illustrates requests made during the work week (the values above the moving average) and on the weekends (the values below the moving average). Note the two lowest points on the chart: 12/25/2004 (Christmas) and 7/10/05 (mostly likely reflects lack of use during the week of the 4th of July). Overall trend shows an increase in usage of *The National Map* over time.

2.4 Most Popular Data Categories

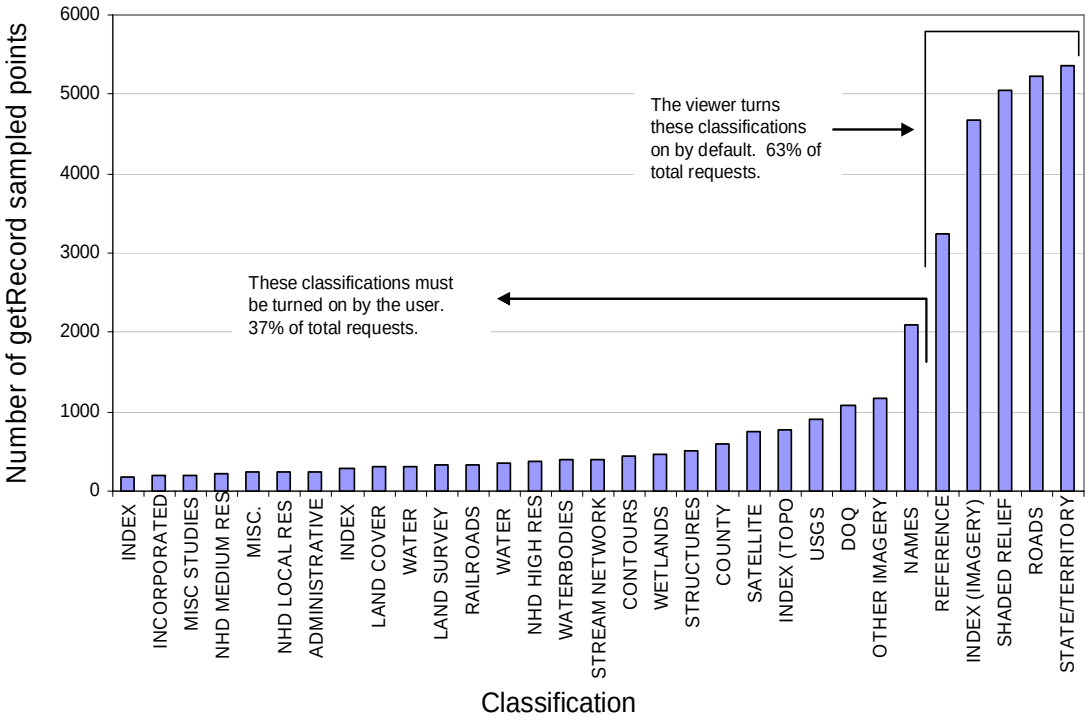


Figure 4. Number of getRecord requests per Classification or Data Category

Figure 4 illustrates that most users use the viewer pretty much as they find it: layers turned on by default tend to stay on, and layers not turned on by default tend to stay off.

2.5 Most Frequent User Destinations by Spatial Extents or Viewscale

Figure 5 illustrates that most users tend to zoom into a fairly small extent to find the information they need.

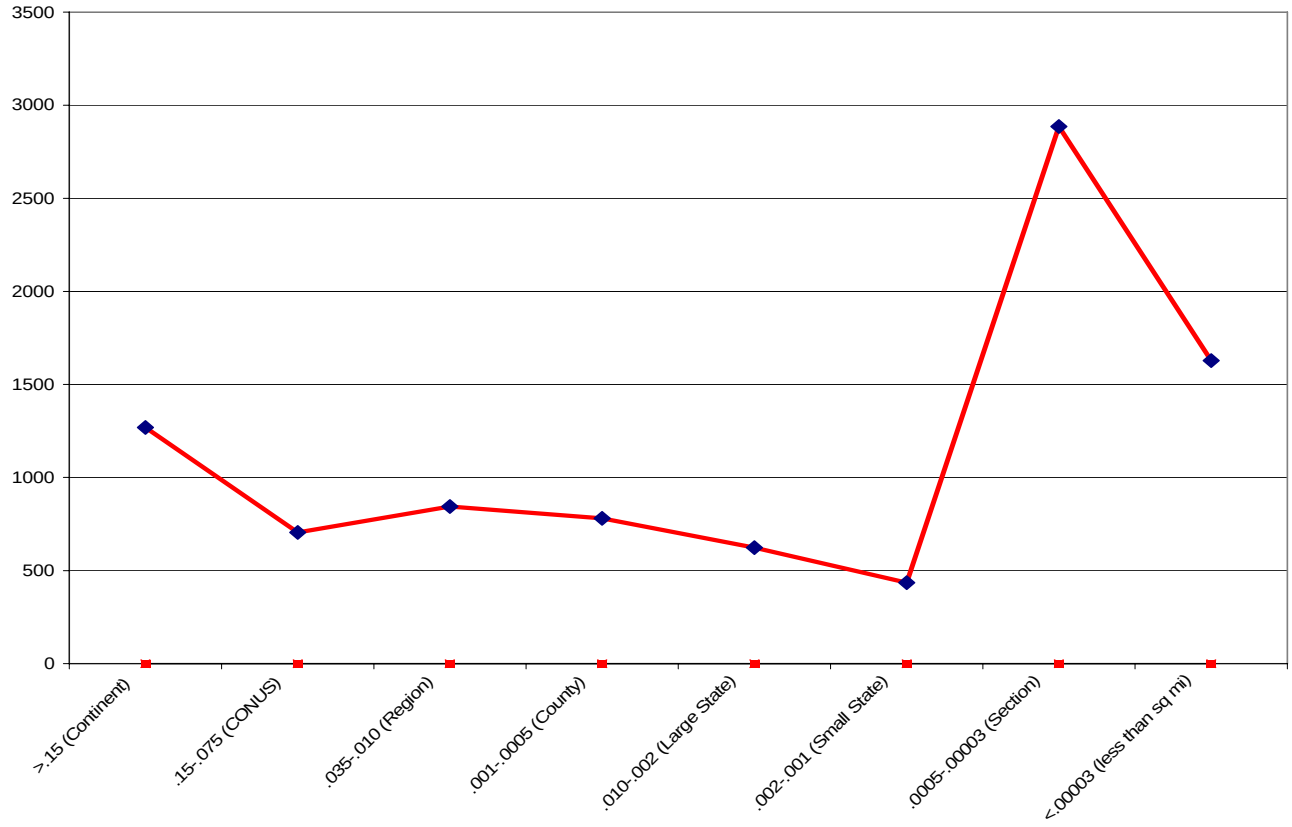


Figure 5. Most common viewscale destinations

By far, the most frequent viewscale used is between .0005 and .00003 decimal degrees per pixel or in a relative scale range of 1:12600 to 1:210000.

3 Discussion

The metrics presented in this paper are relatively narrow. Even so, they provide several reasons to be optimistic about *The National Map*:

- In FY04, the reliability of services that contribute to *The National Map* showed overall improvement. The variation of availability between services decreased. In the last two months of FY04, the median availability of services reached 99%

- In FY05 we saw several service providers make major physical and software changes to their services, resulting in short periods of reduced availability. In the last two months of FY05, the median availability of services was above 98%.

- In FY05, *The National Map* grew at the rate of eight services and 340 layers per month. The "doubling time" of the database in FY05, based on FY04 growth rates, is about 3 months. It is clear that this growth is becoming exponential. [1]
- Though the data holdings are growing fairly rapidly, the resources needed to maintain the system are growing very slowly, or perhaps not growing at all. The catalog support teams (CST) at Mid-Continent and Eastern Region did not grow in FY05. In fact, the resources needed for routine data management of the catalog actually shrank slightly, freeing people to work on training, documentation, and new procedures. This is all evidence that *The National Map's* database-driven design, developed and initially implemented in FY02-03, is robust and scalable.[2]

More and different metrics for evaluating the reliability, growth, and use of *The National Map* are obviously desirable. We do not have a good measure of the overall availability of the public application from an end-user perspective, let alone any measures of customer satisfaction. These are complicated problems that are starting to be addressed in FY06.

4 Additional Reading

The following documents are posted on the catalog web site, http://mcmcweb.er.usgs.gov/catalog/

- For a more complete discussion of how the service checker behaves, see section 7.2 of "Registering Web Map Services in *The National Map* Catalog." Section 7 of this document also discusses security concerns and system loading.
- Three previous reports on service availability contain more detailed data about the early months of service monitoring, and also more detail about the reliability and performance of the service checker software: "Availability of *The National Map* Web Map Service I" , "Availability of *The National Map* Web Map Service II" and *"Availability and Growth of The National Map in Fiscal Year 2004"*.
- "Tutorial Introduction to *The National Map* Catalog" contains an explanation of the overall functioning of *The National Map* system, with emphasis on the role of the catalog database and service

[1] Whether these growth rates are high, low, or insignificant depends on your perspective. Compared to the size and growth rate of data registries populated with automated crawlers (think Google), *The National Map* doesn't even qualify as background noise. On the other hand, the system is already quite large by traditional GIS standards. It also has goals for data integration, partnership building, data maintenance, service reliability, and cartographic quality that are far higher than anything a fully automated system would attempt. These ambitions make smaller size and lower rates of growth necessary and even desirable. The tradeoffs between speed, volume, and quality are severe, and will be a source of continuing debate for the foreseeable future.

[2] The fact that reports like this one can be written at all is another endorsement of the database approach.

5 Additional Usage Data

The following are charts depicting viewer requests by the following breakouts:

- Weekday Usage
- Weekend Usage
- Daily Usage
- Daily Trends

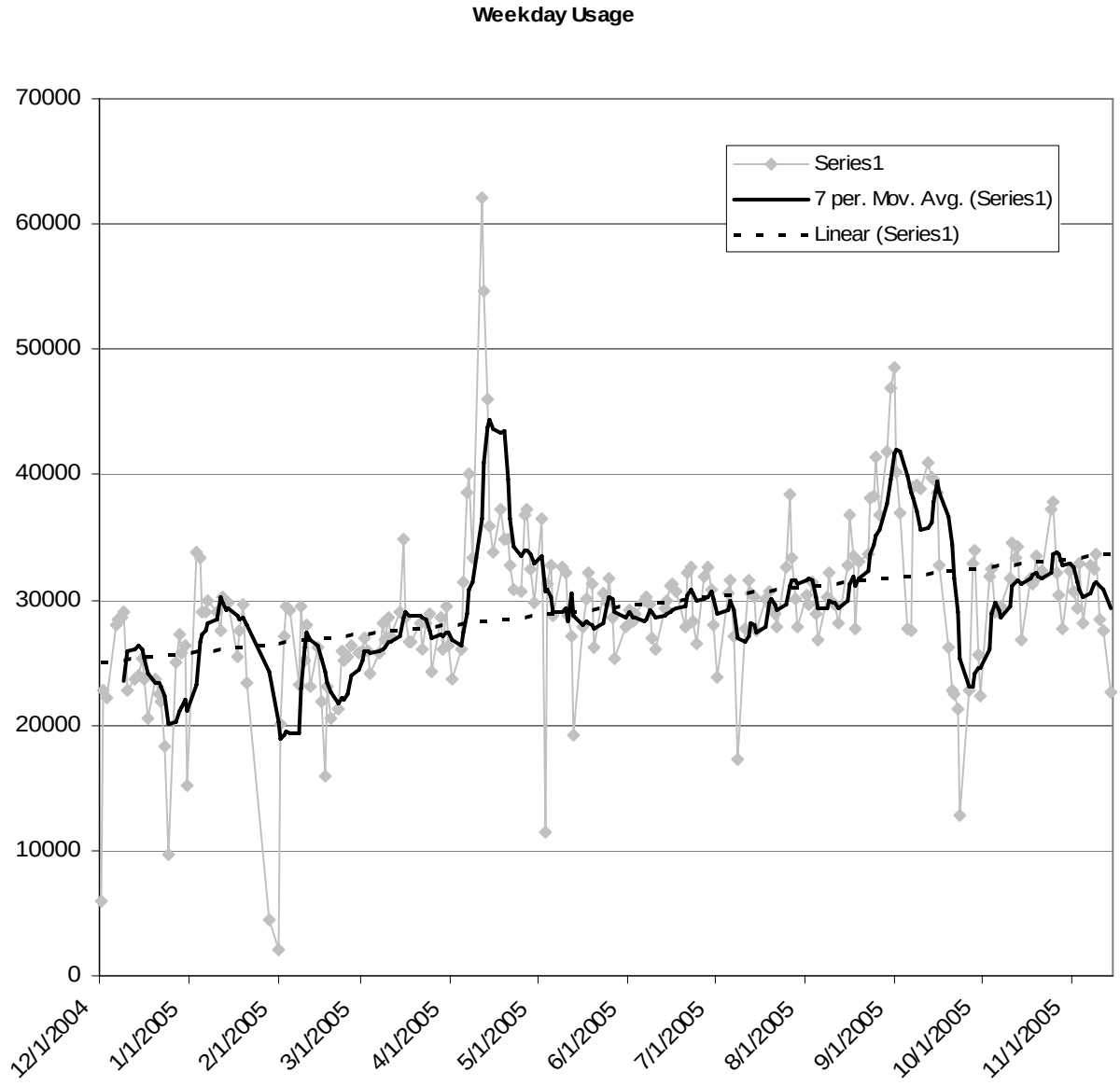


Figure 6. Displays number of viewer screen refresh requests per weekday

From December 1, 2004 until November 14, 2005, the average number of viewer screen refresh requests per weekday has increased from approximately 25,000 requests per day to approximately 32,000 requests per day

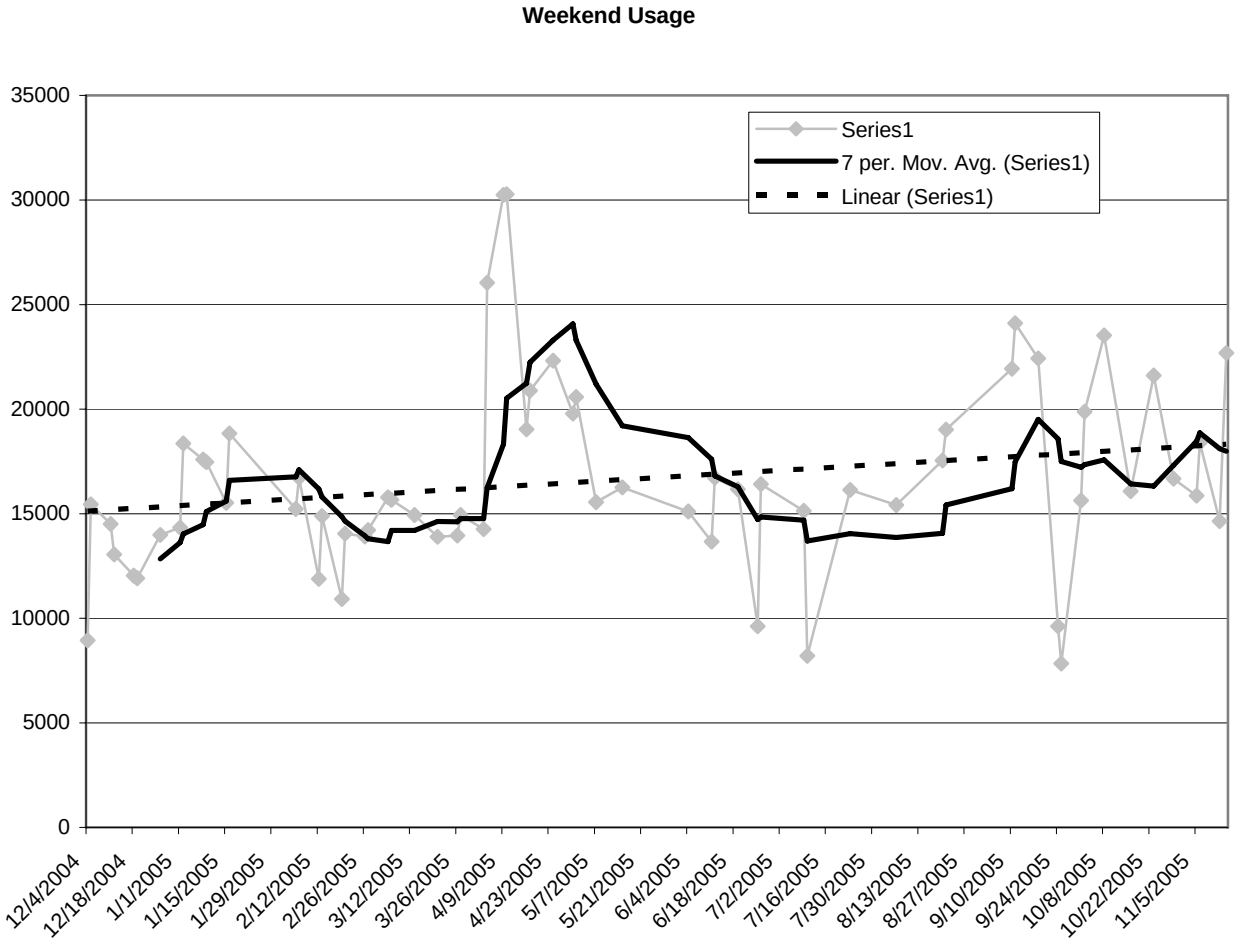


Figure 7. Displays number of viewer screen refresh requests per weekend

From December 1, 2004 until November 14, 2005, the average number of viewer screen refresh requests per weekend has increased from approximately 15,000 requests per day to approximately 17,500 requests per day

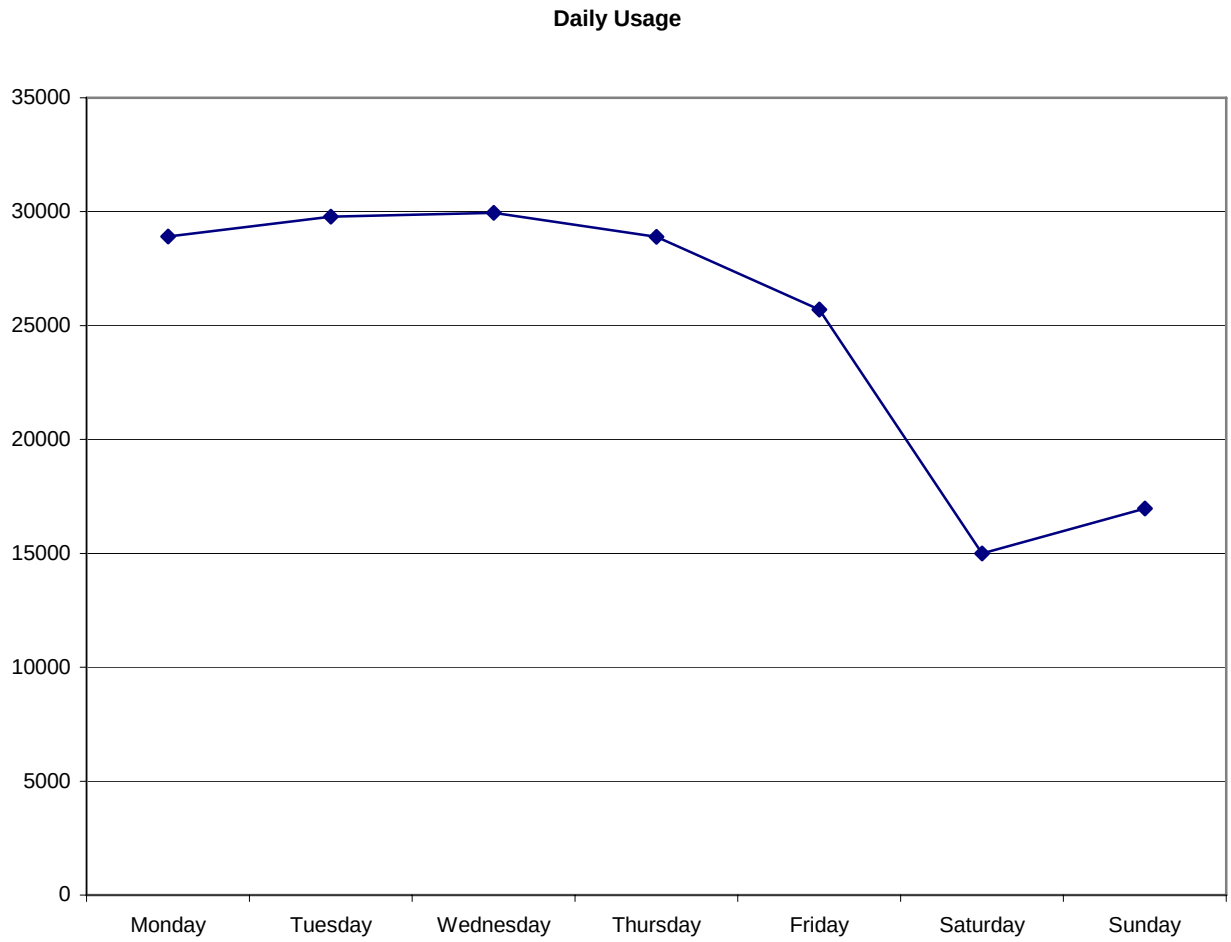


Figure 8. Displays average number of viewer screen refresh requests per day

Figure 8 shows that Tuesday and Wednesday are the two busiest days for *The National Map* viewer in terms of number of viewer screen refresh requests

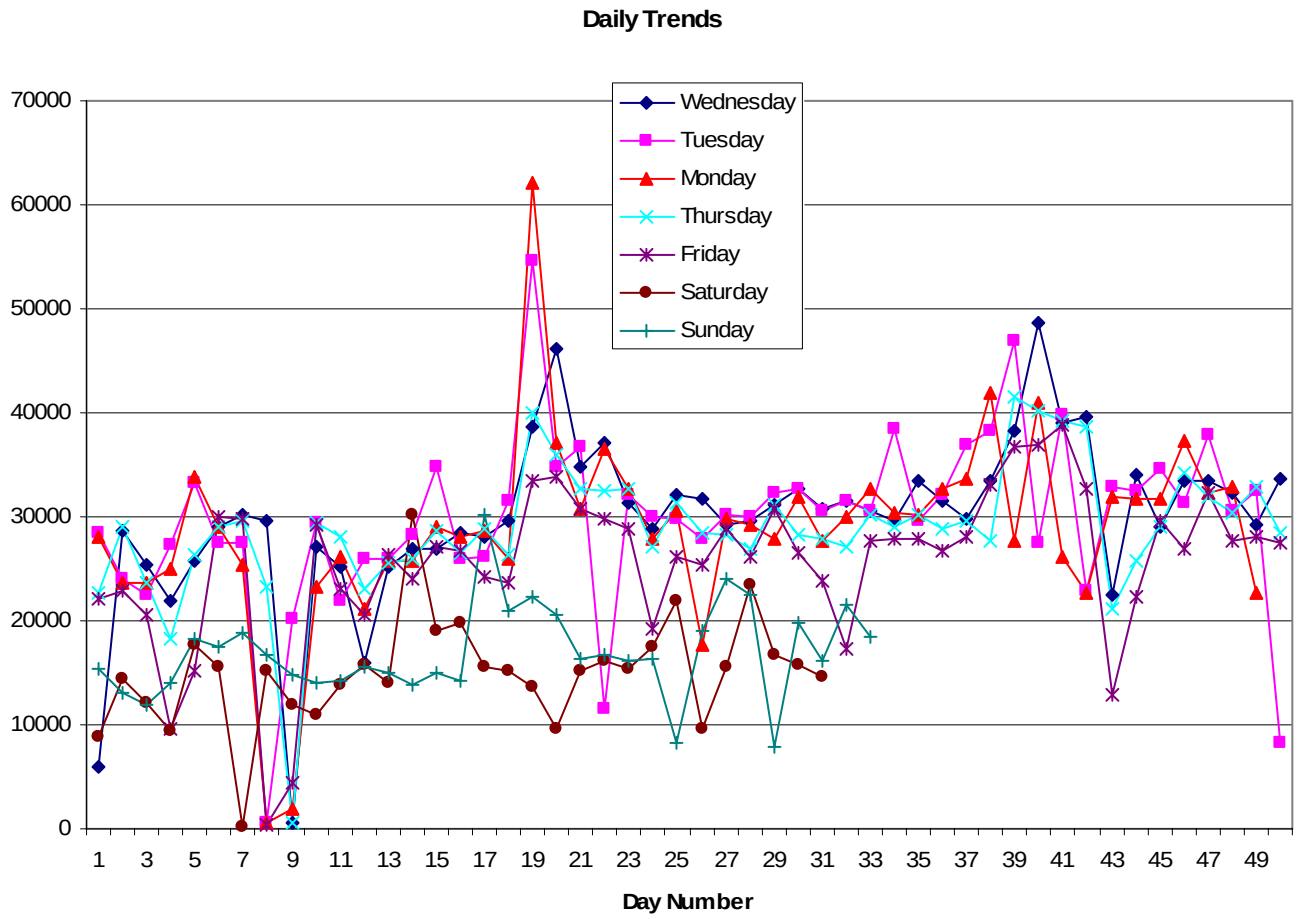


Figure 9. Displays actual number of viewer screen refresh requests per day

6 Measures of Customer Satisfaction

The following is a sampling of the comments received via the catalog email address. These comments are unsolicited and generally reflect a positive acceptance of *The National Map*, but clearly are not enough to obtain any accurate measure of customer satisfaction.

My pleasure! The National Map Viewer is a really good site and makes the investigation on areas and land cover very easy and fun to do. So thank you for making such a brilliant site on the internet!
Best Regards
M S
Houston, Tx 77084

As a first time user to The National Map Viewer I just want to thank you for putting such a great free resource on the web. It's nice to see my tax money being spent on something that's actually useful. As a resident of the Bay Area, I've always been a consumer of the real-time earthquake maps, but this was my first use of your interactive GIS.

And thank you again for making me less regretful of the taxes I pay.
J W

I just wanted to say, as a geography major, that your National Map Viewer is EXCELLENT! and Thanks!
B

Gentlemen and Ladies,
I would like to express my sincere gratitude, thanks, and my awed admiration for the wonderful National Map available on the web. It is a product that more Americans should know about. In my opinion, it is an amazing piece of work and a true national treasure.
From one American citizen and taxpayer to another: Thank you for your fine work.
J R
Chesapeake, VA

USGS Map team,

Thank you for making the online National Map available! I just discovered it recently and am finding it very useful for my work as a biologist.
D R